Hidden Markov Model(HMM)을 이용한 시스템의 이상진단에 관한 연구

조용훈, 이성근, 안대명, 황규석
부산대학교 화학공학과

Fault Diagnosis for System using Hidden Markov Model(HMM)

Yong Hoon Cho, Sung Gun Lee, Dae Myung An, Kyu Suk Hwang
Dept. of Chemical Engineering, Pusan National University

서론

화학공장은 대량생산을 목적으로 수많은 기계장치, 센서, 작동기 제어장치, 전기 장치, 그리고 감시 및 제어시스템 등으로 구성되어 있다. 또한, 안전과 환경에 대한 사회적인 관심과 에너지와 물질의 절약에 대한 요구의 증가로 인해 공정의 운전 조건이 더욱 엄격해지고 매우 복잡한 공정 특성을 나타내고 있다. 이런 복잡한 화학공정의 조업 중에 발생하는 여러 가지 문제를 해결하기 위하여 다양한 이상 진단 시스템 방법론이 사용되고 있다. 이 이상 진단 시스템방법론은 공정에서 측정된 센서 데이터에 의존하고 있으므로 올바른 센서 데이터의 해석은 매우 중요하다. 공정 모니터링과 추론 알고리즘은 센서 데이터를 해석·결정하여 이상 상태를 파악하고 유용한 정보를 추출한다. 그러나 잘못된 시발점의 선택, 알맞은 모델의 부족, 전자파의 방해·간섭 등으로 이런 센서 데이터의 해석에 오류가 발생하기 쉽다. 이 같이 진단 과정에서 불확실한 요소가 첨가된다면 비록 그 결과가 바르게 나왔더라도 우리는 이 결과를 완전히 믿을 수 가 없다. 따라서 진단 과정에서 결과의 불확실함은 또 다른 장애가 된다고 할 수 있다. 그러므로 문제는 신뢰할 수 없는 테스트 결과의 불확실함이 주는 상태가 좋은지/아닌지(good/faulty)를 결정하는 것이다. Hidden Markov Model(HMM)은 시스템의 진단 과정에 도움이 되는 두 가지의 특징이 있다. 첫째, HMM은 확률적 과정을 2중으로 묘사하는 능력이 있다. 시스템 안에서의 이상 상태는 2중으로 된 확률적 과정으로 숨겨져 있기 때문에 직접적으로 관측할 수 없다. 이 부분은 다른 연속적인 테스트 결과의 열들 중에서 가장 최적열을 찾으므로 해결할 수 있다. 둘째, HMM은 시스템 상태의 테스트 결과에서의 순간 확률, 초기 상태 분배와 상태 전이 확률에 의해 묘사되는 Parametric model 이다. 이 HMM의 두 가지 특징은 진단 결과의 불확실함을 해결하는데 많은 도움을 줄 것이다.

본론

1. HMM의 공식화

HMM(Hidden Markov Model) : 통계적 과정을 갖는 유한 상태 구조로 시간에 따른 상태간의 변화를 결정하는 천이 확률과 각 상태에서의 출력확률로 구성된다.

$\boxed{\lambda=(\Pi,\ A,\ B)\ \rightarrow\ \text{Model Parameters}}$

Π : 초기상태 확률분포의 집합

$\Pi=\{\pi_1, \pi_2, \pi_3, \cdots, \pi_N\}$ π_i는 초기상태가 s_i 인 확률, N : 상태(state) 수

$\pi_i=P(s=s_i \mid \lambda)$, $1\leq i \leq N$

A : 상태천이 확률분포의 집합

$A=\{a_{ij}\}$, $1\leq i, j\leq N$ $a_{ij}=P(s_{t+1}=s_j \mid s_t=s_i)$, $1\leq t\leq T$

a_{ij}는 시간 t에서 상태 s_i로부터 시간 t+1에서 s_j로의 천이확률

여기서, $a_{ij}\geq 0$, $\sum_{j=1}^{N} a_{ij}=1$ 인 조건을 갖는다.

B : 상태에서의 관측확률의 집합

$B=\{b_j(k)\}$, $1\leq j\leq N$, $1\leq t\leq T$

$b_j(k)=P(o_t=v_t \mid s_t=s_j)$ $b_j(k)$는 시간 t에서 상태 j에서의 심볼 k의 관측확률

여기서, $1\leq k\leq M$: M은 상태당 관측심볼의 수

2. 문제 해결

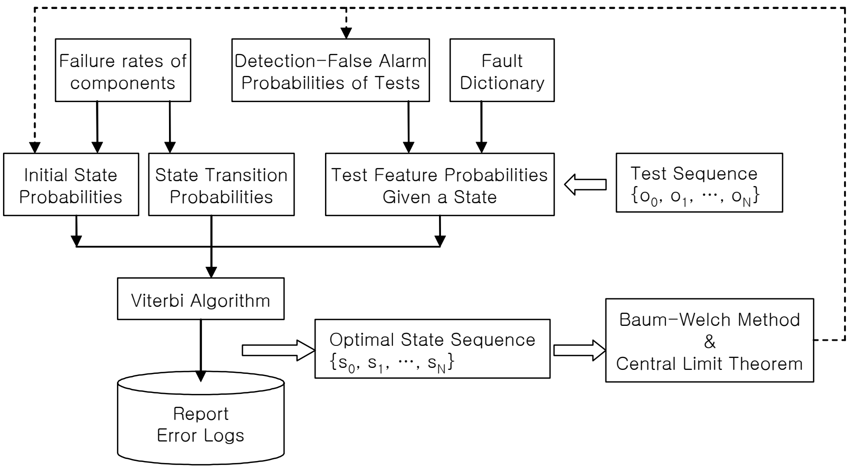

Fig. 1. Overview of Fault diagnosis using HMM

HMM을 사용하여 이상진단문제를 푸는데 다음과 같은 2가지 특별한 문제가 요구된다.
1)시스템 상태의 가장 최적의 상태열을 찾기위한 해석 문제(Decoding problem).
2)관측열에 대하여 확률을 최대화하는 모델 변수를 예측하는 문제(Estimation problem).
이상의 기본적인 문제점에 대한 해결 방안으로 전향-후향 알고리즘, viterbi 알고리즘, Baum-Welch 재추정 알고리즘이 널리 사용되고 있다.

A. 전향 알고리즘(forward algorithm)과 후향 알고리즘(backward algorithm)

전향변수 α_t는 초기상태에서 시작하여 $o_1, o_2, \cdots, o_t$를 생성하면서 상태 i에 도달하는 확률로 정의한다. 이 전향확률은 다음과 같은 재귀적으로 계산하는 것이 가능하다.

$\alpha_1(i)=\pi_i b_i(o_1)$, $1\leq i\leq N$

$\alpha_{t+1}(j)=\left[\sum_{i=1}^{N} \alpha_t(i)a_{ij}\right]b_j(o_{t+1})$, $1\leq t\leq T-1$

후향변수 β_t는 최종상태에서 처음으로 o_T, …, o_{t+2}, o_{t+1}을 생성하면서 상태 i에 도달하는 확률로 정의하면

$\beta_T(i)=1$, $1 \le i \le N$

$\beta_t(i)= \sum_{j=1}^{N} a_{ij} b_j(o_{t-1}) \beta_{t+1}(j)$, $t=T-1, T-2, \cdots, 1$

이런 전향변수 와 후향변수는 재추정(restimation)시 사용된다.

B. 재추정(restimation)

전방향 확률과 역방향 확률을 이용하여 학습패턴에 대하여 모델의 특정상태 사이에서 발생한 전이 횟수의 기대값 $\xi_t(i, j)$와 특정 상태로부터 발생한 전이 횟수의 기댓값$\gamma_t(i)$를 구한다.

$\xi_t(i, j)=P(s_t=i, s_{t+1}=j \mid o,\lambda)$

$$= \frac{\alpha_t a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{P(o \mid \lambda)} = \frac{\alpha_t a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_t a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}$$

$\gamma_t(i)= \sum_{j=1}^{N} \xi t(i,j)$

C. 변경(update)

주어진 학습 패턴에 대해 전이 횟수의 기댓값 $\xi_t(i, j)$와 $\gamma_t(i)$를 이용하여 확률 매개변수 $\lambda=(\pi, A, B)$ 는 다음과 같은 새로운 값으로 된다.

$\bar{\pi}_1=\gamma_1(i)$, $1 \le i \le N$

$$\bar{a}_{ij} = \frac{\text{상태}i\text{로부터상태}j\text{로의전이횟수}}{\text{상태}i\text{로부터의전이횟수}} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}$$

$$\bar{b}_j(k) = \frac{\text{상태}j\text{에서심볼 }v_k\text{를관측한횟수}}{\text{상태}j\text{에서심볼의관측횟수}} = \frac{\sum_{t=1,\ s.t.\ o_t=v_k}^{N} \gamma_t(i)}{\sum_{t=1}^{N} \gamma_t(j)}$$

D. Viterbi Decoding

viterbi 알고리즘은 최적 원리에 입각한 동적 프로그래밍 기법의 하나이다. 이 알고리즘은 임의 상태에 이르는 경로 비용 또는 확률을 계산할 때, 이전 상태까지의 비용과 이전 상태들로부터 현재 상태로의 전이 비용을 곱하는 방식으로 순환계산하는 기법이다. 이 점은 마르코프 모델의 가정과 그에 따르는 시간적 제약에 밀접한 관계를 가지므로 마르코프 모델내의 최적 경로를 찾는데 적용할 수 있다. viterbi 알고리즘은 초기화, 순환계산, 종료의 단계로 수행된다.

3. HMM 구현시 고려사항

HMM을 실제 응용분야에서 적용할 때에는 다음과 같은 몇 가지 고려 해야할 사항들이 있다.

A. 다중 관측열(Multiple Observation Sequence)

단일 관측열을 이용하여 HMM의 parameter들을 재추정 하기에는 불충분하다. 따라서 좋은 모델을 얻기 위해서는 다중 관측열로부터 parameter를 재추정 해야한다. 이것은 단일 관측열에 적용되는 HMM을 다중 관측열에 적합하도록 사용 할 수 있다.

B. Multiple Codebook

여러 개의 코드북을 사용하면 양자화 에러가 적다는 점에 착안하여 HMM의 parameter를 추정할 수 있다.

C. Scaling문제

컴퓨터 시뮬레이션시 underflow가 발생하지 않도록 하기 위하여 이 방법을 사용한다. 이 방법은 전향변수와 후향변수가 시간 $1 \le t \le T$인 동안에도 컴퓨터의 유효 범위 안에서 계산될 수 있도록 scaling 계수를 곱하는 것이다. 그러나 이러한 모든 비율계수는 계산이 종료되는 시점에서는 Baum-Welch 알고리즘의 정확도를 보장하기 위하여 모두 제거되어야 한다.

D. Smoothing

이것은 출력 심볼이 관측열에 나타나지 않을 시 사용하는데 학습 관측열을 증가시키거나 상대적으로 HMM의 크기를 감소시켜 자유매개변수의 수를 작게 설정하는 parameter typing방법과 전체 학습 데이터에 대해 이를 분할한 또 다른 parameter에 가중치를 두어 삽입하는 deleted interpolation이 있다.

결론

이 연구에서 두 가지에 중점을 두었다. 관측된 테스트 결과를 기초로 HMM parameter들을 유도하였고, 특정한 실재적인 데이터에 의해 HMM의 거동을 분석하였다. 그 방법은 테스트 결과들과 시스템 이상 상태들을 고정된 확률적 과정으로 모델링하여 결합시키는 것이다. 이 결합은 하나의 고정된 HMM Frame-Work를 사용하여 동적인 이상진단을 모델링하는 것을 두가지의 독립된 관점으로 연구 할 수 있게 도와 줄 것이다. 이 연구에 제시 된 기술로 인해 우리는 다중 이상 문제(multiple fault problem)를 해결할 수 있다. Baum-Welch 알고리즘에 의한 Model Parameter들은 진단과정의 구성을 학습하는데 도움을 줄 것이고 Viterbi 기술은 평가의 복잡함을 해결해 줄 것이다.

참고문헌

1. Jie Ying, T. Krishna R. Pattipati, “A Hidden Markov Model-Based Algorithm for Fault Diagnosis with Partial and Imperfect Tests”.
2. Warren J. Ewens, Gregory R. Grant, “Statistical Methods in An Introduction.” 327-348
3. 김상운 “식별 알고리즘을 중심으로 한 패턴인식 입문” 137-152